



HAL
open science

a cross-entropy based multiagent approach for multimodal activity chain modeling and simulation

Tai-Yu Ma, Jean-Patrick Lebacque

► **To cite this version:**

Tai-Yu Ma, Jean-Patrick Lebacque. a cross-entropy based multiagent approach for multimodal activity chain modeling and simulation. 2009. halshs-00310903v2

HAL Id: halshs-00310903

<https://shs.hal.science/halshs-00310903v2>

Preprint submitted on 15 Jun 2009 (v2), last revised 11 May 2011 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A CROSS-ENTROPY BASED MULTIAGENT APPROACH FOR MULTIMODAL ACTIVITY CHAIN MODELING AND SIMULATION

Tai-Yu Ma¹, Jean-Patrick Lebacque²

¹tai-yu.ma@let.ish-lyon.cnrs.fr, Laboratoire d'Economie des Transport (LET), France

²lebacque@inrets.fr, Institut National de Recherche sur les Transports et leur Sécurité (INRETS),
France

1 INTRODUCTION

Modeling activity chaining behavior is a complicated research theme in activity-based travel demand analysis. Basically, an activity chain denotes an ordered sequence of planned/unplanned activities connected by trips on a daily basis under space-time, institutional and household constraints. The realization of an activity program involves heuristical rescheduling behavior at the moment of execution. This rescheduling process is difficult to observe and the resulting travel choice usually depends on individual's available time and related constraints. Previous studies indicated that travel decision with respect to the realization of an activity program is opportunistically undertaken depending not only on the activities previously conducted but also on the opportunities provided by later occurring activities (Doherty et al., 1998; Lee et al., 2002). The traditional trip-based traffic demand analysis cannot catch the dependency of sequential travel decision on the basis of the realization of the activity chain.

This paper suggests a dynamic activity chaining model and proposes an arc-based Cross Entropy (CE) approach to approximate activity-chain-based dynamic user equilibrium (Rubinstein, 1999; Helvik and Wittner, 2001; Ma and Lebacque, 2007b). A multiagent approach is proposed for complex activity chaining behavior modeling and simulation on a multimodal road network. Different with traditional trip-based approaches, the proposed activity-based approach considers travel demand as subsidiary decisions resulting from activity chaining behavior, provided individual's (agent's) scheduled activity program. Based on the average performance of travel/activity choice updated on a day-to-day basis, these agents are guided to utilize more attractive destination/departure time/path choice. In this study, an arc-based CE method is proposed by extending previous path-based CE method. The complex activity-chain-based dynamical user equilibrium condition is formulated by applying dynamic programming approach. Each agent is considered as a rational expected utility maximizing user aiming to, at each activity location, maximizing total expected utility gain for not yet implemented activities in his/her activity program. This optimal strategy for travel/activity choice is based on the principle of optimality of Bellman, resulting in an approximate of activity-chain-based dynamic user equilibrium.

2 MODEL FORMULATION

We model and simulate agent's travel/activity demand in a sequential way, provided his/her pre-planned activity program. The agent's activity chaining behavior is stated as follows. Let each agent have a fixed activity program to engage in a period of time. Keeping planned activities in mind, each agent makes his/her transport mode and destination-chain choice at home and then conducts sequentially the departure time and route choice for the next activity to be conducted. We assume that each activity possesses a gross economic value/utility. As the activity value is generally difficult to measure, its value is assumed to have an exponential distribution (Leurent, 1999). The travel/activity choice probability depends on its average performance measured by the gross economic value of activity minus its generalized travel cost. Although the proposed model simplifies complex agent's decision making, it captures the essential of travel-cost/activity-value tradeoff in the activity program realization process.

2.1 Activity program and gross value of activity chain

Let an activity program g be defined as an ordered sequence of planned activities that one agent would like to engage during a period of time, i.e.:

$$g = \{0, 1, 2, \dots, i, \dots, n\} \quad (1)$$

where i denotes i th planned activity. Each activity is characterized by its desired starting time τ and its activity type σ . The sequence of desired starting times of activities satisfy $0 \leq \tau_0 < \dots < \tau_n$. A trip-chain of activity program g with origin o is defined as an ordered sequence of paths linking all scheduled activity locations:

$$u = \{r_1, r_2, \dots, r_n\}, \quad \forall u \in U_g^o, \quad \forall r_i \in R_{k_{i-1,i}} \quad (2)$$

where o is home location, and r_i is the path connecting two consecutive activities $i-1$ and i . U_g^o denotes the set of trip-chains for activity program g with departure origin o and $R_{k_{i-1,i}}$ the set of paths connecting the origin-destination (OD) pair k of activity $i-1$ and i . Let the network be represented by a directed graph $G = (V, E)$ with V being the set of nodes and E the set of arcs. We denote O the set of origins, D the set of destinations of activities and $K = \{(i, j) \mid \forall (i, j) \in \{O \cup D\}\}$ be the set of possible OD pairs. Following previous study (Ma and Lebacque, 2006), we assume that the activities are distributed at destinations with limited supplies. Its economic value/utility with respect to the destination d follows an exponential probability distribution function (pdf) as:

$$P_d(v) = \lambda_d \exp(-\lambda_d(v - m_d)), \quad \forall d \in D \quad (3)$$

where $v \geq m_d$ and $\lambda_d \geq 0$.

By assuming the additivity of economic value, the net activity value obtained at destination is calculated as its gross activity value minus its generalized path cost. The later is composed of two parts, one being its trip time and the other the early/late arrival penalty with respect to expected starting time of activity. Based on boundedly-rational behavior assumption (Mahmassani and Chang, 1985), an indifference interval is introduced to reflect the tolerable schedule variation. The generalized trip cost of the path r_i , connecting the destinations of the scheduled activity $i-1$ and i , is formulated as:

$$C_{r_i}(t) = \pi_{r_i}(t) \times \mu + \mu_\alpha \times \max(0, \tau_i - \Delta - t_{d_i}^{\text{arr}}) + \mu_\beta \times \max(0, t_{d_i}^{\text{arr}} - \tau_i - \Delta) \quad (4)$$

where $\pi_{r_i}(t)$ is the travel time of path r_i when entering the initial point of the path at time t , $t_{d_i}^{\text{arr}}$ the arrival time at the destination d_i of activity i . μ is the unitary economic value of travel time. μ_α and μ_β are unitary penalty associated with early and late arrival, respectively.

Δ is the half of tolerable schedule delay interval without penalty.

The net value of activity i obtained by an agent at destination depends on the agent's: (1) leaving time from previous activity $i-1$; (2) arrival time to next activity i ; (3) path choice. The net activity value is defined as:

$$v_i^*(r_i, t_i, t_{i-1}) = v_i(t_i) - C_{r_i}(t_{i-1}) \quad (5)$$

where $v_i(t_i)$ is the economic value of activity i obtained at destination when arriving at time t_i . $C_{r_i}(t_{i-1})$ is the generalized travel cost of path r_i when entering the initial point of the path at time t_{i-1} , i.e. the time of leaving previous activity $i-1$.

Based the above equation, the net activity value of activity chain $v_{og}(u, t)$ is calculated as the sum of net activity values of each activity, i.e.

$$v_{og}(u, t) = \sum_{\forall i \in g} v_i^*(r_i, t_i, t_{i-1}), \quad \forall g, \quad \forall o \quad (6)$$

where u denotes a trip chain. t is the entering time of transport network.

2.2 Dynamical user equilibrium based on activity chain

We consider a long-term predictive dynamic user equilibrium given agents' fixed activity programs. Different with traditional trip-based models, the proposed activity-chain-based model implies that agent's optimal decision rule is based on the maximization of future expected net activity value gains under traffic and activity availability dynamics. The dynamic user equilibrium problem in question is essential the same as a sequential decision making problem for which each agent tends to find a non-cooperative optimal rule for limited resources. This problem has been recognized as difficult to solve and been remaining an

active research issue in the domain of artificial intelligence, computer science and economics (Littman, 1996).

Hence, the activity-chain-based dynamical user equilibrium can be stated as an extension of Wardrop principle (Wardrop, 1952) as:

“Agents of the same attributes settings, i.e. the same origin, mode availability and activity program, the expected total net activity value at each decision making stage, resulting from destination, departure time and path choice, is equal and no less than that of unused choice alternatives”.

For agents at activity $i-1$, the optimal travel choice for the next activity i is based on the principle of optimality of Bellman:

$$a_i^*(t_{i-1}) = \arg \max_{a_i \in A(t_{i-1})} E[\sum_{s \geq i} v_s^*(a_s(t_{s-1})) | s_{s-1}(t_{s-1})] \quad (7)$$

where $a_i^*(t_{i-1}) \stackrel{def}{=} (r_i^*, t_i, t_{i-1})$ is the optimal travel choice at time t_{i-1} for next activity i with respect to route choice r_i^* , time t_{i-1} at the activity $i-1$ and time t_i arriving to the activity i , conditional on agent's current state $s_{i-1}(t_{i-1})$. $A(t_{i-1})$ is the feasible travel choice set at time t_{i-1} with respect to departure time and path choice. $v_s^*(a_s(t_{s-1}))$ denotes the net activity value of next activity s . Given the above optimal decision rule, the activity-chain-based dynamic user equilibrium is the sequence of optimal solution of the equation (7). The difficulty remains on the computing of optimal probability distributions with respect to agent's destination-chain, departure time and path choice. We address this issue in the next section.

3 SOLUTION ALGORITHM

To solve user equilibrium problem under multimodal simulation-based traffic flow model, we propose the arc-based CE approach to derive iteratively optimal probability distributions guiding agents to construct their routes incrementally to destinations. This proposed approach is very suitable for multiagent framework since it is arc/node-based for agent's path choice, avoiding the difficulty of path enumeration for all OD pairs. The CE approach derives optimal choice probability distribution based on average performance sampled at previous day (iteration) such that agents of the same class (competing the same activity/transport supply) shift to more valuable alternatives at next iteration. It is also shown that this method is more efficient than dynamical system approach since it optimizes the adjustment step size towards unique/multiple equilibrium points. The reader is referred to (Ma and Lebacque, 2007b; Ma, 2007c) for more details and applications.

Note that the path choice probability in the proposed approach is the multiplication of sequential arc/node choice probability of its path. The rationale of this approach is that the sequence of probability laws should converge towards a law concentrated on optimal paths.

The detail of proposed algorithm for solving dynamic activity-chain-based user equilibrium problem is described as follows.

Main algorithm

The proposed approach to compute activity-chain-based dynamic user equilibrium is stated as follows. Each agent constructs a path-chain connecting each of activity destinations of his/her activity program. Each path connecting two consecutive activity destinations is constructed as follows. Conditional with the destination chain determined at origin, an agent selects departure time for next activity at current location, and then incrementally moves to an next outgoing arc/node until the destination. When arriving at destination, the obtained gross activity value and general travel cost is evaluated. This process is repeated until the agent's activity program is accomplished. After all agents complete his/her program, the related choice probability distributions with respect to destination-chain, departure time and outgoing arc/node choice are updated according to the performance of related travel choice.

Step 1: Initialization

Initialize uniform pdf for destination-chain and departure time choice for all agents. The departure time choice interval for any activity i is considered within the desired starting time interval $[\tau_{i-1}, \tau_i]$ of two consecutive activities i and $i-1$. For outgoing node choice at node x with respect to OD pair k , we associate a set of OD dependent probability distribution P_{xk} , $\forall k \in K$ for all nodes. The initialization of P_{xk} is stated as follows (Rubinstein, 1999).

1. Initialize uniform pdf for outgoing node choice for all $x \in V$ and $k \in K$.
2. Generate N random paths for all $k \in K$ based on $P_{xk} = \{p_{yxk} \mid y \in \Lambda^+(x)\}$, $\forall p_{yxk} \geq 0$ with $\Lambda^+(x)$ being the set of outgoing nodes of x . Evaluate path costs and order them from lowest to highest. Then calculate ρ -quantile of the path cost $C_{[\rho N]}$, i.e. $[\rho N]$ th lowest cost ($[a]$ denotes the smallest integer greater than or equal to a). Note that ρ is set small, say $0.1 \leq \rho \leq 0.3$, such that each outgoing node choice probability is not too small.
3. Update p_{yxk} as:

$$p_{yxk} = \frac{\sum_{r \in R_k} 1_{[C_r \leq C_{[\rho N]}]}(x, y)}{\sum_{r \in R_k} 1_{[C_r \leq C_{[\rho N]}]}(x)}, \forall y \in \Lambda^+(x) \quad (8)$$

where $1_{[C_r \leq C_{[\rho N]}]}(x, y)$ is an indicator if (x, y) belongs to the path r where its OD travel cost satisfies $C_r \leq C_{[\rho N]}$. Increase iteration by 1.

4. If maximum iteration is achieved or P_{xk} keeps unchanged, then stop, otherwise go to 2.

Step 2: Route construction and net activity value evaluation

Each agent moves from origin to each of activity destinations according to the destination-chain, departure time and outgoing node choice. When arrival to each of destination, the agent obtains a gross economic value and decides the departure time for next activity. The suitable departure time choice interval T_i for next scheduled activity i is considered between agent's arrival time t_{i-1} at activity $i-1$, and the desired starting time τ_i of next activity i . We assume that agents choose a departure time interval and then randomly choose a departure time instant within it. This process is repeated until all activities in his/her activity program are completed.

Step 3: Choice probability update

The choice probability distributions with respect to destination-chain, departure time and outgoing node choice are updated as follows. Note that the choice probability distributions are updated when all agents have finished his/her activity program.

Destination chain choice probability update

The choice probability of destination-chain \mathbf{d} , conditional on agent's activity program g and his/her origin o . The choice probability is updated based on the summation of normalized net activity values obtained at each of activity destinations. It is defined as:

$$p_{\mathbf{d}^s \mathbf{og}}^{w+1} = p_{\mathbf{d}^s \mathbf{og}}^w \frac{e^{-\tilde{v}_{\mathbf{d}^s \mathbf{og}}^w / \gamma_{\mathbf{og}}^w}}{\sum_{\mathbf{d}' \in \mathbf{D}_{\mathbf{og}}} p_{\mathbf{d}' \mathbf{og}}^w e^{-\tilde{v}_{\mathbf{d}' \mathbf{og}}^w / \gamma_{\mathbf{og}}^w}}, \quad \forall g, \mathbf{d} \quad (9)$$

with $\tilde{v}_{\mathbf{d}^s \mathbf{og}}^w = \frac{v_{\mathbf{d}^s \mathbf{og}}^{w, \max} - \hat{v}_{\mathbf{d}^s \mathbf{og}}^w}{v_{\mathbf{d}^s \mathbf{og}}^{w, \max}}$, where $\hat{v}_{\mathbf{d}^s \mathbf{og}}^w = \frac{\bar{v}_{\mathbf{d}^s \mathbf{og}}^w}{\bar{v}_{\mathbf{og}}^w}$ with $\bar{v}_{\mathbf{d}^s \mathbf{og}}^w$ being the average net activity value for

agents of the same destination-chain choice with origin o and activity program g , similarly for

$\bar{v}_{\mathbf{og}}^w$. $v_{\mathbf{d}^s \mathbf{og}}^{w, \max} = \max \{ \hat{v}_{\mathbf{d}^s \mathbf{og}}^w \mid \forall \mathbf{d} \in \mathbf{D}_{\mathbf{og}} \}$, $\mathbf{D}_{\mathbf{og}}$ is the set of possible destination-chains for origin o

and activity program g . $\gamma_{\mathbf{og}}^w$ is the control parameter with respect to o and g at the iteration w .

The above equation represents agent's normalized relative net activity value with respect to the destination chain \mathbf{d} , origin o and activity program g .

Departure time choice probability update

The probability of discretized departure time choice h , conditional on agent's activity program g , OD pair k , and desired starting time τ of next activity i , is updated as follows:

$$p_{h^s k g \tau_i}^{w+1} = p_{h^s k g \tau_i}^w \frac{e^{-\tilde{v}_{h^s k g \tau_i}^w / \gamma_{k g \tau_i}^w}}{\sum_{h' \in H_{\tau_i}} p_{h' k g \tau_i}^w e^{-\tilde{v}_{h' k g \tau_i}^w / \gamma_{k g \tau_i}^w}}, \quad \forall k, g, \tau, h \quad (10)$$

where $\tilde{v}_{hkg\tau_i}^w = \frac{v_{hkg\tau_i}^{w,\max} - \hat{v}_{hkg\tau_i}^w}{v_{kg\tau_i}^{w,\max}}$ is the relative normalized net value of activities with $\hat{v}_{hkg\tau_i}^w = \frac{\bar{v}_{hkg\tau_i}^w}{\bar{v}_{kg\tau_i}^w}$. $v_{kg\tau_i}^{w,\max} = \max \{ \hat{v}_{hkg\tau_i}^w \mid \forall h \in H_{T_{\tau_i}} \}$ denotes the maximum of normalized net activity value for agents of the same attributes in terms of k , g , and τ_i . $H_{T_{\tau_i}}$ is the set of the discretized time intervals within the time interval $T_{\tau_i} = [t_{i-1}, \max(t_{i-1}, \tau_i)]$, $\forall i$. Similarly, the average net activity value $\bar{v}_{hkg\tau}^w$ and $\bar{v}_{kg\tau}^w$ are defined as that for precedent step.

Outgoing arc/node choice probability update

The choice probability of next outgoing node y from node x , conditional on agent's OD pair k , is updated as follows:

$$p_{yxk}^{w+1} = p_{yxk}^w \frac{e^{-\tilde{C}_{yxk}^w / \gamma_{xk}^w}}{\sum_{y' \in \Lambda^+(x)} p_{y'xk}^w e^{-\tilde{C}_{y'xk}^w / \gamma_{xk}^w}}, \forall x \in V, \forall k \in K \quad (11)$$

where $\tilde{C}_{yxk}^w = \bar{C}_{yxk}^w / \bar{C}_{xk}^w$ is the normalized average travel cost of OD pair k . \bar{C}_{yxk}^w and \bar{C}_{xk}^w , represent agents' realized average travel cost for paths passing the arc (x, y) and the node x , respectively.

Step 4: Stopping criteria

When maximum iteration is achieved or the derived choice probability distributions are stable, stop, otherwise go to step 2.

4 NUMERICAL STUDY

In this section, we present numerical results for a grid road network of 60 nodes and 208 arcs. The length of arcs (road section) is randomly set between 1km and 5km. The number of lanes is set uniformly as 2 for all arcs. For the demand setting, two classes of users are considered: (1) car with travel speed of 90 km/hr; (2) bus with that of 45km/hr. For each class, we set 200 agents at origin nodes 11 and 50 (Fig. 1). The activity destinations are located at node 7, 14, 26, 28, 35, 42, and 56. The parameter m_d (average gross activity value at destination d) of Eq. (3) at the above destinations is set as 10, 20, 40, 30, 25, 15, and 55 (euros), respectively. λ_d is set as 4 for all destinations. For simplicity, each agent has an activity program of two activities with desired starting time at 9:00 and 11:00, respectively. The activity supply is set as sufficiently large (4000 activities/destination) for all destinations. The traffic flow model is based on the extended point queue model (Kuwahara and Akamatsu, 1997; Ma and Lebacque, 2007a). However, more elaborated models could be applied in real application (Lebacque et al., 2007). The discretized departure time choice interval is set as 5 minutes. The unitary

travel cost and early/late arrival penalty are set as 7 euros/hour, 4 euros/hour and 15 euros/hour, respectively.

Fig. 2 (left). shows the convergence result of total net activity value of agents with respect to origins and transport mode. The difference of total net activity value between 4 classes of agents (2 modes multiplied by 2 origins) reflects that agents of class car and origin 50 have more advantage to occupy better activities (lower travel time to arrive to the destinations with higher activity value). Fig. 3 depicts the evolution of choice probability of destination-chain 26-56 for agents from origin 11 (left) and 56-26 for origin 50 (right). The results show that the probabilities of best destination-chain choice (with higher gross activity values (40 and 55 euros/activity, respectively) than other destinations) increase rapidly to 1. However, the probabilities of other destination-chain choices decrease rapidly to 0. Fig. 4 shows the evolution of choice probability over outgoing nodes along the shortest path 11-12-13-14-15-16-26. The results show that the choice probability of outgoing node on the shortest path increase gradually (Fig. 4) and the general travel cost decreased rapidly at early stage and then stabilize (Fig. 2 (right)).

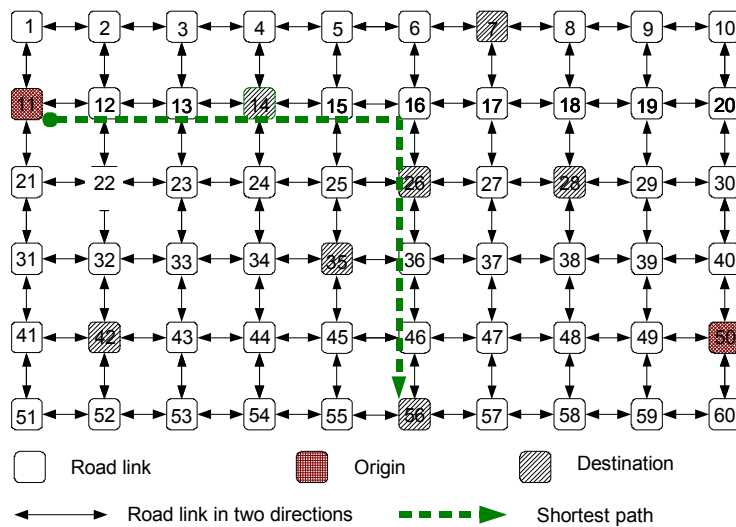


Figure 1. Grid road network example

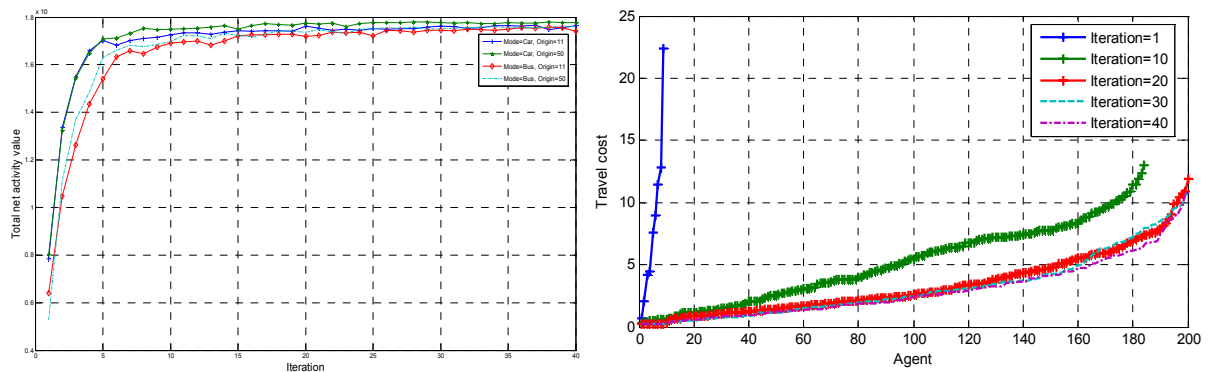


Figure 2. Convergence of all agents' obtained net activity value based on activity program realization (Left) and the evolution of ordered travel cost of agents from activity 1 (node 26) to activity 2 (node 56)

5 CONCLUSION AND DISCUSSIONS

In this study, a dynamical activity-chaining model is proposed to capture individual's travel behavior under temporal and spatial constraints. The resulting dynamic system equilibrium is approximated on the basis of agent's daily learning and adaptation for its activity chaining according to the overall performance of the activity program execution. We have proposed an arc-based Cross Entropy approach to compute the approximate of activity chain based user equilibrium in a multimodal network. The solution algorithm is based on the concept that the equilibrium choice probability distributions concerning destination-chain, departure time and path choice is iteratively estimated based on the expectation of travel/activity choice performance. Future extensions include the study of the behavior of proposed CE approach based on a multimodal transport network and also the equilibrium stability as the presence of multiple equilibriums.

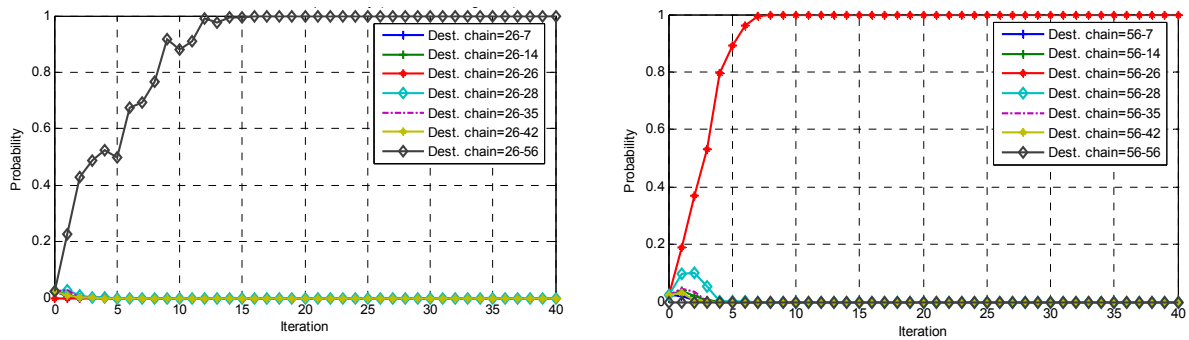
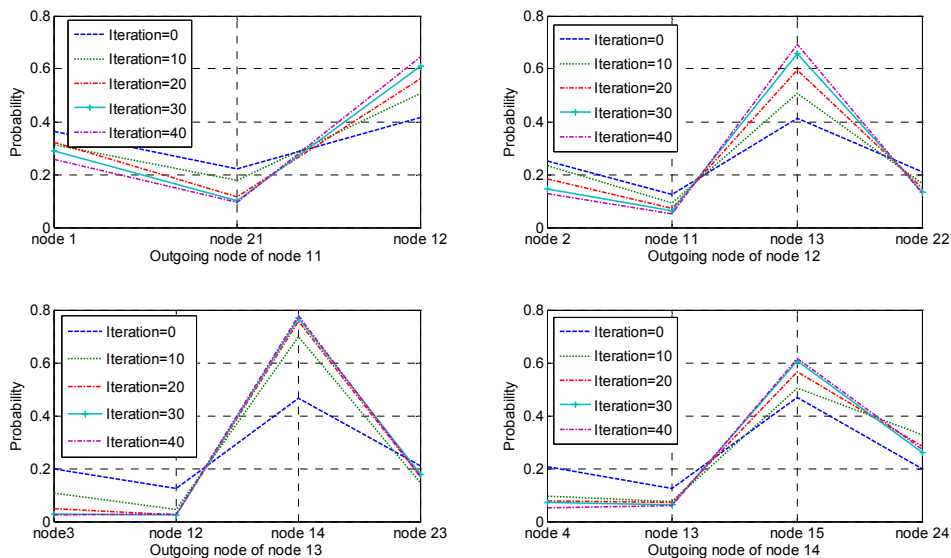


Figure 3. Evolution of destination-chain choice probability. (Left: agents from origin 11 with car. Right: agents from origin 50 with car)



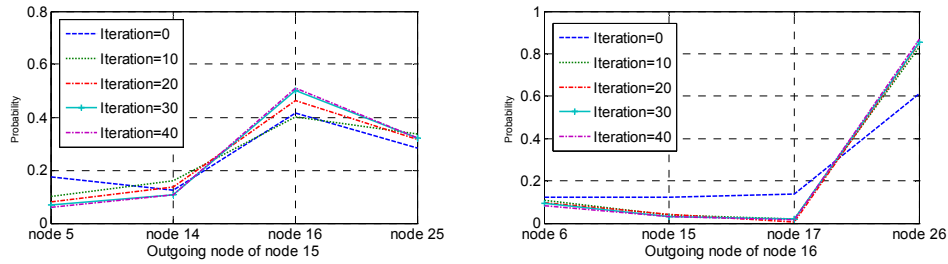


Figure 4. Evolution of outgoing node choice probability on a shortest path (mode=car, origin=11, destination=26)

REFERENCE

1. Doherty, S.T., Miller, E.J., Axhausen, K.W., Gärling, T., (1998). A conceptual model of the weekly household activity-travel scheduling process. *Proc. of the Network on European Communications and Transport Activities Research*, Euroconference, Israel.
2. Helvik, B.E., Wittner, O., (2001). Using the Cross-Entropy Method to Guide/Govern Mobile Agent's Path Finding in Networks. *Proc. of the 3rd International Workshop on mobile Agents for Telecommunication Applications*.
3. Kuwahara M. Akamatsu, T. (1997). Decomposition of the reactive assignments with queues for many-to-many origin-destination pattern. *Transportation Research B*, 31(1), pp. 1-10.
4. Lebacque J.P., Mammari, S., Haj-Salem, H. (2007). Generic second order traffic flow modelling. *Proc. of the 17th International Symposium on Transportation and Traffic Theory (ISTTT)*, London.
5. Lee, M.S., Chung, J.-H., McNally, M.G. (2002). An Empirical Investigation of the Underlying Behavioral Processes of Trip Chaining. Center for Activity Systems Analysis. Paper ITS-UCI-AS-WP-02-6.
6. Leurent, F. (1999) Accessibility to Vacant Activities: a Novel Model of Destination Choice. *Proc. of the European Transport Conference, Seminar F*, pp. 307-318, PTRC, London.
7. Littman, M.L. (1996). *Algorithms for sequential decision making*. Thesis. Brown University.
8. Ma, T.Y., Lebacque, J.P. (2006). Modeling activity choice distribution and road choice behaviour on a network with side constraints. *Proc. of the 11th Euro Working Group on Transportation*.
9. Ma, T.Y., Lebacque, J.P. (2007a). A multi-agent approach to dynamic traffic assignment based on activity. *Proc. of the sixth Triennial Symposium on Transportation Analysis (TRISTAN VI)*, Phuket Island, Thailand.
10. Ma, T.Y., Lebacque, J.P. (2007b). A cross entropy based multi-agent approach to traffic assignment problems. *Proc. of the Traffic and Granular Flow*, Orsay, France.

11. Ma T.Y. (2007c) *Modèle dynamique de transport basé sur les activités*. PhD dissertation, ENPC (Ecole Nationale des Ponts et Chaussées).
12. Mahmassani, H. S., Chang, G.-L. (1985). Dynamic aspects of departure time choice behaviour in a commuting system: Theoretical framework and experimental analysis. *Transportation Research Record*, No. 1037, pp. 88-101.
13. Rubinstein, R.Y., (1999). The Cross-Entropy Method for Combinatorial and Continuous Optimization. *Methodology and Computing in Applied Probability*, 1(2), pp. 127-190.
14. Wardrop, J.G. (1952). Some Theoretical Aspects of Road Traffic Research. In: *Proc. of the Institution of Civil Engineers, Part II*, vol. 1, pp. 325-362.